

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L1	0	(identif\$5 and features and duplicate and document and scores and reciprocal and fragments and words).clm.	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2005/07/07 10:12

clm 09/713,733

Interference Search History.

12 Phrasier: a system for interactive document retrieval using keyphrases

Steve Jones, Mark S. Staveley

August 1999 **Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval**

Full text available:  pdf(625.73 KB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

Keywords: evaluation, interactive retrieval interface, keyphrase-based retrieval, query interface


13 Identifying topics by position

Chin-Yew Lin, Eduard Hovy

March 1997 **Proceedings of the fifth conference on Applied natural language processing**

Full text available:  pdf(996.73 KB)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)



 [Publisher Site](#)

This paper addresses the problem of identifying likely topics of texts by their position in the text. It describes the automated training and evaluation of an Optimal Position Policy, a method of locating the likely positions of topic-bearing sentences based on genre-specific regularities of discourse structure. This method can be used in applications such as information retrieval, routing, and text summarization.

14 Using suffix arrays to compute term frequency and document frequency for all substrings in a corpus

Mikio Yamamoto, Kenneth W. Church

March 2001 **Computational Linguistics**, Volume 27 Issue 1

Full text available:  pdf(1.80 MB) 

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)


[Publisher Site](#)

Bigrams and trigrams are commonly used in statistical natural language processing; this paper will describe techniques for working with much longer n -grams. Suffix arrays (Manber and Myers 1990) were first introduced to compute the frequency and location of a substring (n -gram) in a sequence (corpus) of length N . To compute frequencies over all $N(N + 1)/2$ substrings in a corpus, the substrings are grouped into a manageable number of equivalence classes. In this way, a ...

15 How reliable are the results of large-scale information retrieval experiments?

Justin Zobel

August 1998 **Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval**

Full text available:  pdf(1.16 MB)

Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

16 Document filtering with inference networks

Jamie Callan

August 1996 **Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval**

Full text available:  pdf(809.41 KB)

Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)



USPTO

[Subscribe \(Full Service\)](#) [Register \(Limited Service, Free\)](#) [Login](#)
Search: ☒ The ACM Digital Library ☐ The Guide

+identifying, +duplicate, +documents, +scores



THE ACM DIGITAL LIBRARY


[Feedback](#) [Report a problem](#) [Satisfaction survey](#)

Published before November 2001

Found 201 of 119,420

Terms used identifying duplicate documents scoresSort results
by

relevance

Display
results

expanded form

☒ Save results to a Binder☒ Search Tips☐ Open results in a new
window

Try an Advanced Search

Try this search in [The ACM Guide](#)

Results 1 - 20 of 200

Result page: [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [next](#)

Best 200 shown

Relevance scale ☐ ☐ ☐ ☐ ☐

1 [Experiments on incorporating syntactic processing of user queries into a document retrieval strategy](#)

A. F. Smeaton, C. J. van Rijsbergen

May 1988 **Proceedings of the 11th annual international ACM SIGIR conference on Research and development in information retrieval**

Full text available: pdf(1.94 MB)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

Traditional information has relied on the extensive use of statistical parameters in the implementation of retrieval strategies. This paper sets out to investigate whether linguistic processes can be used as part of a document retrieval strategy. This is done by predefining a level of syntactic analysis of user queries only, to be used as part of the retrieval process. A large series of experiments on an experimental test collection are reported which use a parser for noun phrases as part o ...

2 [Mechanized Indexing Methods and Their Testing](#)

John O'Connor

October 1964 **Journal of the ACM (JACM)**, Volume 11 Issue 4

Full text available: pdf(1.04 MB)

Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

3 [Creating and evaluating multi-document sentence extract summaries](#)

Jade Goldstein, Vibhu Mittal, Jaime Carbonell, Jamie Callan

November 2000 **Proceedings of the ninth international conference on Information and knowledge management**

Full text available: pdf(186.89 KB)

Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

4 [Efficient text summarization using lexical chains](#)

H. Gregory Silber, Kathleen F. McCoy

January 2000 **Proceedings of the 5th international conference on Intelligent user interfaces**

Full text available: pdf(586.39 KB)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)